QRISK®3 Cardiovascular Disease Risk

Search Definition

2023

ClinRisk Ltd

Revision History

Revision date	Document Version	Summary of Changes
04/01/2023	V0.1	Initial version

Contents

1 Purpose of this document

This document defines the search criteria needed to calculate a QRISK®3 10-year cardiovascular disease score from general practice computer systems in the UK. It also gives details of the table structure required for the QRISK®3 Batch Processor Utility.

The variable definitions in this document are also useful for system suppliers who intend to use the ClinRisk Ltd software development kit to embed the score within their own applications/software.

Once the individual patient level data have been extracted from a GP clinical computer system, the resulting CSV file can be processed using the QRISK®3 'batch processor' to generate a QRISK®3 10-year score on every patient, using default values where patients do not have systolic blood pressure, body mass index or cholesterol recorded.

For an implementation of the QRISK®3 see qrisk.org/three.

2 Search definition

This section defines the patient inclusion criteria, the input parameters and their definitions.

2.1 Patient inclusion criteria

The QRISK®3 algorithm should be applied to patients who meet all of the following criteria.

- Currently registered on the search date
- Aged between 25 and 84 inclusive on the search date
- Free from a diagnosis of cardiovascular disease (SNOMED-CT code group 16779) prior to the search date.

2.2 Missing values and 'out of range' data

- We don't expect all patients on GP clinical computer systems will have complete data for all the clinical values needed to generate an actual QRISK3® score.
- For patients with complete data the QRISK3® software will generate an actual QRISK3® score.
- For patients where one or more item is missing the QRISK3® will generate an estimated score.

The assumptions regarding missing data as follows

- If the patient does not have a recorded family history of premature coronary heart disease in a first degree relative under the age of 60 they are assumed not to have a family history. This should be coded as such in the input file or function call.
- Patients who do not have smoking status recorded are assumed to be non-smokers. This should be coded in the input file or function call.
- Patients who have a postcode which is not in the postcode database are assumed to have a Townsend score of zero (this is the mean value of Townsend score nationally). If the postcode is not known, then the cell should be left blank and the batch processor will substitute a zero. If both Townsend score and postcode are given, the Townsend score supplied is used.
- If the patient does not have a recorded value for body mass index, systolic blood pressure, or cholesterol ratio, then the software will calculate an estimated value.

The following rules are applied automatically by the software to data which are 'out of range' or which are missing. The values are approximately equal to the 1st and 99th centile of values derived from the population of patients on the QResearch database.

Table 1: rules for out of range clinical values

	Acceptable/Credible Range	Action if values outside of credible range
Systolic blood pressure (mm Hg)	70 to 210	<70 is set to 70 >210 is set to 210
Body mass index (kg/m^2)	18 to 47	<18 is set to 18 >47 is set to 47
Cholesterol/HDL ratio	1 to 12	<1 is set to 1 >12 is set to 12

Valid Townsend scores range between -8.0 and + 12.0

2.3 Input parameters

- Table 2 defines the input parameters needed to calculate the QRISK®2 cardiovascular disease 10 risk score. It also shows the table structure needed for the Batch processor.
- The input table needs to contain one row per record with each field separated by a comma (comma separated file).
- The table needs to follow this exact format including order of the column, field types and field names.
- The Code group ID refers to the identity of the cluster of codes used in the search definition. The contents of each SNOMED-CT code cluster can be found in the accompanying SQL database file.

Table 2: Input parameters needed to calculate a QRISK3® score

Full Field name	DataType	English description	Code group ID	Coding of variable
row_id	Integer	Unique row identifier	n/a	n/a
age	Integer	Patients age in years calculated on the search date	n/a	n/a
sex	Char	Patients sex	n/a	M=males F = females
b_cvd	Boolean	Binary variable to denote diagnosis of CVD recorded at any time prior to the search date	16779	1 = yes, 0 = no
b_AF	Boolean	Atrial fibrillation at any time prior to the search date	24	1 = yes, 0 = no
b_atypicalantipsy	Boolean	Second generation 'atypical' antipsychotic - prescribed two or more issues in the previous 6 months (includes amisulpride, aripiprazole, clozapine, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, sertindole, or zotepine)	1169	1 = yes, 0 = no
b_corticosteroids	Boolean	Systemic corticosteroids – prescribed two or more issues in the previous 6 months	1194	1 = yes, 0 = no
b_migraine	Boolean	Diagnosis of migraine prior to the search date	39	1 = yes, 0 = no
b_ra	Boolean	Rheumatoid arthritis at any	68	1 = yes, 0 = no

		time prior to the search date		
b_renal	Boolean	Chronic renal disease at any time prior to the search date	322	1 = yes, 0 = no
b_semi	Boolean	Diagnosis of severe mental illness (psychosis, severe depression, manic depression, schizophrenia) at any time prior to the search date	3187	1 = yes, 0 = no
b_sle	Boolean	Diagnosis of systemic lupus erythematosis at any time prior to the search date	70	1 = yes, 0 = no
b_treatedhyp	Integer	SNOMED-CT coded diagnosis of hypertension (group 21) at any time in the patient's records AND On antihypertensive treatment if 1+ script in last 6 months prior to the search date. Antihypertensive treatment includes the following • Thiazides (group 740 or BNF 2.2.1) • Beta blockers (group 751 or BNF 2.4) • ACE inhibitors (group 759 or BNF 2.5.5.1) • A2A (group 760 or BNF 2.5.5.2) • Calcium Channel Blockers (group 765 or BNF 2.6.2) Patients should only be included if both criteria are satisfied i.e. on treatment and have a diagnosis of hypertension. BNF chapters can be used instead	21 740 751 759 760 765 Or use BNF groups for drugs	1 = yes 0 = no
bmi	Double (nullable)	The most recently recorded body mass index for the patient recorded prior to the search date recorded in the last 5 years.	200	n/a
diabetes_cat	integer	categorical variable to denote no diabetes, type 1 diabetes or type 2 diabetes diagnosed at any time prior to the search date and	n/a	0 = no diabetes 1 = type 1 diabetes 2 = type2 diabetes

		recorded in the patients record. See DiabetesCategoryMapping in the accompanying code lists (1913 for Type 1 and 2411 for Type 2)		
ethnicity	Integer	Ethnic group in 17 categories.	n/a	0 =not recorded 1 =British 2 =Irish 3 =Other White Background 4 =White & Black Caribbean 5 =White & Black African 6 =White & Asian 7 =other mixed 8 =Indian 9 =Pakistani 10= Bangladeshi 11 =Other Asian 12 =Caribbean 13 =Black African 14 =Other Black 15 =Chinese 16 Other ethnic group 17 Not stated
fh_cvd	Boolean	Family history of coronary heart disease in a first degree relative under the age of 60 years recorded before the search date	404	1 = yes, 0 = no
rati	Double (nullable)	The most recent ratio of total serum cholesterol/HDL recorded in the last 5 years. The ratio is either directly directly recorded (eg using group 405) or can be calculated using group 16 divided by group 367	16 367 405	n/a
smoke_cat	Integer (nullable)	The most recently recorded smoking status 5 levels i.e. • non-smoker (group	2239 2238	0= non smoker 1 =ex-smoker

		2239) • Ex-smoker(group 2238) • Light-smoker(group 2244) • Moderate-smoker(group 2242) • Heavy-smoker (group 2243)	2244 2242 2243	2 =light smoker 3= moderate smoker 4= heavy smoker Null if missing
sbp	Integer (nullable)	The most recent systolic blood pressure prior to search date recorded in the last 5 years	198	n/a
sbps5	Double (nullable)	Standard deviation of all of the available systolic blood pressures prior to the search date recorded in the last 5 years	198	n/a
town	Double (nullable)	The Townsend score associated with the output area of a patient's postcode based on the 2001 census data.	n/a	n/a
postcode	String	Patient's postcode. This can be missing if Townsend score is supplied.	n/a	n/a

3 Appendix

3.1 Smoking status

Smoking status is categorised into 5 groups. This is the same as the categorisation used for smoking status in other scores e.g QStroke® and QDiabetes®.

Smoking information can either be recorded in the medical record as a SNOMED-CT Code or the actual number of cigarettes can be recorded or both.

Step 1: identify most recent SNOMED-CT code from the table below recorded in the last 5 years prior to the search date

Step 2: if the most recent code is associated with a numeric_value for the number of cigs/day then categorise patients based on the number of cigs/day as follows:

0 = non-smoker
1 = ex-smoker
2 = current light smoker (<10 cigs/day)
3 = current moderate smoker (10-19 cigs/day)
4 = current heavy smoker (20+ cigs/day)

Step 3: If the most recent SNOMED-CT Code is not associated with a numeric_value for the number of cigs/day, then categorise patients according to the SNOMED-CT code using the mapping in the table above.

Step 4: there will be a small percentage of patients who are current smokers according to the SNOMED-CT code but who do not have a number of cigs/day recorded. These should be coded as moderate smokers so an estimated score can be calculated.

Note the following:

- This categorisation has one group for ex-smokers rather than trying to distinguish between ex light, ex moderate and ex heavy smokers.
- This categorisation does not distinguish between cigarette/cigar and pipe smokers
- this could be a complex bit of coding as some patients will have multiple codes recorded on the same day and the codes and units can conflict
- for example, a patient might have 5 cigs/day coded under a SNOMED-CT code which allows units to be recorded but then have a SNOMED-CT code indicating they are a heavy smoker. If this occurs, then the units should be used in preference to the SNOMED-CT code even if the two codes occur on the same dat.
- Alternatively, a patient might be coded as a non-smoker but then have the number of cigs recorded in the numeric_value field. In this circumstance, you should take the most recent entry and base the categorisation on that.
- In summary, it is important that this procedure is followed in sequence and that you check that all patients are correctly coded.

3.2 Ethnicity

- The categorisation of ethnicity is the same as that used for other QPrediction scores including QStroke® and QDiabetes®
- The mapping of SNOMED-CT codes to each category can be found in accompanying code lists.

3.3 Diabetes status

- QRISK3 distinguishes between type 1 and type 2 diabetes.
- Some patients with diabetes will have SNOMED-CT codes for type 1 and type2 both recorded in their medical record. It is likely that the most recently recorded code is the most reliable.
- Therefore, to categorise patients into type 1 or type 2, then identify the latest SNOMED-CT code recorded in the medical record. If it is Type 1, then code the patient as type 1. If it is type 2, then code the patient as Type 2.
- If a patient has two codes (one for type 1 and one for type2) recorded on the same day and that is the latest recorded entry, then assume the patient is Type2.
- The final categorisation should ensure that patients are EITHER Type 1 or Type 2 or neither.

4 References

Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017;357 doi: 10.1136/bmj.j2099